MIT Technology Review
Insights

Produced in association with

MEDIDATA
acornai

Researchers are using analytics and existing patient data to ease recruitment, reduce costs, and accelerate timelines.

Clinical trials are better, faster, cheaper with big data

Key takeaways

1. Researchers today can use artificial intelligence and data analytics to speed up the clinical trial process. Instead of having to recruit patients for a traditional control arm – the group that doesn't get the experimental treatment given to the test group – investigators are building "external control arms," which reuse data on control-group patients from past clinical trials.

2. External control arms yield several benefits. They can reduce or eliminate the time normally needed to recruit control patients, expediting access to experimental treatment for patients in the test group. They cut the costs of recruiting control group patients and tracking them during the trial. And using external controls makes it easier to recruit potential participants, because everyone recruited will get the treatment.

3. The US Food and Drug Administration looks favorably on external control arms in general, especially in single-arm trials (a type of trial in which a regular control group is impractical). Replacing traditional control arms with external data faces more scrutiny – but a hybrid design, in which external controls supplement a recruited control arm, is currently under review by the FDA.

Clinical trials have never been more in the public eye than in the past year, as the world watched the development of vaccines against covid-19, the disease at the center of the 2020 coronavirus pandemic. Discussions of study phases, efficacy, and side effects dominated the news. The most distinctive feature of the vaccine trials was their speed. Because the vaccines are meant for universal distribution, the study population is, basically, everyone. That unique feature means that recruiting enough people for the trials has not been the obstacle that it commonly is.

"One of the most difficult parts of my job is enrolling patients into studies," says Nicholas Borys, chief medical officer for Lawrenceville, N.J., biotechnology company Celsion, which develops next-generation chemotherapy and immunotherapy agents for liver and ovarian cancers and certain types of brain tumors. Borys estimates that fewer than 10% of cancer patients are enrolled in clinical trials. "If we could get that up to 20% or 30%, we probably could have had several cancers conquered by now."

Clinical trials test new drugs, devices, and procedures to determine whether they're safe and effective before they're approved for general use. But the path from study design to approval is long, winding, and expensive. Today, researchers are using artificial intelligence and advanced data analytics to speed up the process, reduce costs, and get effective treatments more swiftly to those who need them. And they're tapping into an underused but rapidly growing resource: data on patients from past trials.

Building external controls

Clinical trials usually involve at least two groups, or "arms": a test or experimental arm that receives the treatment under investigation, and a control arm that doesn't. A control arm may receive no treatment at all, a placebo or the current standard of care for the disease being treated, depending on what type of treatment is being studied and what it's being compared with under the study protocol.

It's easy to see the recruitment problem for investigators studying therapies for cancer and other deadly diseases: patients with a life-threatening condition need help now. While they might be willing to take a risk on a new treatment, "the last thing they want is to be randomized to a control arm," Borys says. Combine that reluctance with the need to recruit patients who have relatively rare diseases – for example, a form of breast cancer characterized by a specific genetic marker – and the time to recruit enough people can stretch out for months, or even years. Nine out of 10 clinical trials worldwide – not

External control arms can make clinical trials less expensive by reducing the number of patients that need active management.

just for cancer but for all types of conditions – can't recruit enough people within their target timeframes. Some trials fail altogether for lack of enough participants.

What if researchers didn't need to recruit a control group at all and could offer the experimental treatment to everyone who agreed to be in the study? Celsion is exploring such an approach with New York-headquartered Medidata, which provides management software and electronic data capture for more than half of the world's clinical trials, serving most major pharmaceutical and medical device companies, as well as academic medical centers. Acquired by French software company Dassault Systèmes in 2019, Medidata has compiled an enormous "big data" resource: detailed information from more than 23,000 trials and nearly 7 million patients going back about 10 years.

The idea is to reuse data from patients in past trials to create "external control arms." These groups serve the same function as traditional control arms, but they can be used in settings where a control group is difficult to recruit: for extremely rare diseases, for example, or conditions such as cancer, which are imminently life-threatening. They can also be used effectively for "single-arm" trials, which make a control group impractical: for example, to

Why use an external control arm?

Under certain circumstances, researchers might create a control group using patient data from past clinical trials instead of recruiting a new control group. They might need to do the following:

Run single-arm trials

Find patients for studies in which traditional control groups are not practical—for an implanted device, for example.

Conduct a trial for a rare disease

Sometimes there are few patients with the condition under study. In other cases, patients with life-threatening diseases want to be sure they'll receive the experimental treatment rather than an existing treatment option.

Do rapid preliminary trials

Evaluate whether a treatment is effective enough to pursue in a full clinical trial.

Expedite access to medical treatment

An external control arm can cut recruiting time dramatically.

Source: MIT Technology Review Insights; based on information from the US Library of Medicine

measure the effectiveness of an implanted device or a surgical procedure. Perhaps their most valuable immediate use is for doing rapid preliminary trials, to evaluate whether a treatment is worth pursuing to the point of a full clinical trial.

Medidata uses artificial intelligence to plumb its database and find patients who served as controls in past trials of treatments for a certain condition to create its proprietary version of external control arms. "We can carefully select these historical patients and match the current-day experimental arm with the historical trial data," says Arnaub Chatterjee, senior vice president for products, Acorn AI at Medidata. (Acorn AI is Medidata's data and analytics division.) The trials and the patients are matched for the objectives of the study – the so-called endpoints, such as reduced mortality or how long patients remain cancer-free – and for other aspects of the study designs, such as the type of data collected at the beginning of the study and along the way.

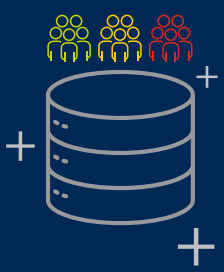
When creating an external control arm, "We do everything we can to mimic an ideal randomized controlled trial," says Ruthie Davi, vice president of data science, Acorn AI at Medidata. The first step is to search the database for possible control arm candidates using the key eligibility criteria from the investigational trial: for example, the type of cancer, the key features of the disease and how advanced it is, and whether it's the patient's first time being treated. It's essentially the same process used to select control patients in a standard clinical trial – except data recorded at the beginning of the past trial, rather than the current one, is used to determine eligibility, Davi says. "We are finding historical patients who would qualify for the trial if they existed today." Once this basic screening stage is completed, statistical matching and weighting techniques are used to match the possible control patients with the patients in the test arm.

External control arms can make clinical trials less expensive by reducing the number of patients that need active management. Borys estimates that each cancer patient costs Celsion tens of thousands of dollars to enroll in a trial and follow throughout the entire protocol. And using external controls may make a study more appealing to potential participants, expediting recruitment.

How external controls work

To build an external control arm, researchers mine a database of past clinical trials to find control subjects who are close matches to the patients in the treatment arm. The information available on the subjects in the database lets researchers match subjects using relevant criteria, such as age, gender, weight, type of condition, and marital status.

Researchers can use several methods to design a clinical trial with an external control arm. Here's one:



Database of past clinical trials
Researchers can enhance the recruited control group with data from patients who were in control groups in past trials for other treatments of the same condition—or they can use the data to build the group from scratch.

TRIAL
Researchers compare the results of the two groups.

TREATMENT ARM
Group receives an experimental treatment for a certain condition.

CONTROL ARM
Group receives no treatment for the condition or a placebo or current standard of care for the condition.

Source: US National Library of Medicine

New insights from "real world" data

Real-world data – the information about all of us that's in our electronic health records, our pharmacy's prescription database, our insurance claims – has the potential to further refine the clinical trial process, make it faster, cheaper, and more accurate, and extend it to post-market surveillance to verify whether new medications and procedures fulfill the promise of their clinical trials.

Medidata is exploring such possibilities with San Francisco health tech company Datavant, which works with health systems and electronic-claims clearinghouses to collect and link data on patients – while preserving their privacy as required by federal law – so that researchers can study the health information that's gathered over time on a single individual, or do advanced analytics on a group of patients that share certain characteristics.

"We're trying to link data from those clinical trial cohorts to the rest of their real-world data, but in a privacy-preserving way," says Jason LaBonte, chief strategy officer at Datavant. "So, if you want to understand more about the patients that were in the trial, you're not stuck if you didn't collect the data."

Datavant works with about 400 health systems and other providers and has access to a vast trove of insurance claim data. LaBonte estimates that among its various data sources, the company has at least some health and other information on about 300 million people in the United States.

Health systems, companies, and organizations that work with Datavant run its software on their patient databases to create a de-identified version of the data. The set of data that identifies someone as a unique individual is replaced with a "linking token." Each institution's data is encrypted so that no other institution can identify individual patients.

Datavant's secret sauce is a complex method of identifying the tokens for the same patient from multiple sources. In that way, a researcher doesn't know the identity of a subject but does know all the subject's diagnoses and treatments over time: across multiple physicians, hospitals, pharmacies, labs, and even insurers. The researcher can track what happens to an individual clinical trial participant after the trial is over – a type of long-term follow-up that's typically expensive and not always practical.

"If I have a subject in a clinical trial, and I want to connect all of her lab records and her electronic medical record, and her insurance claims, data linking via tokenization allows us to link all that data," says Arnaub Chatterjee of Medidata Acorn AI. "We can pull together all those records that exist out in the ether." Clinical trial patients need to give explicit consent for their data to be used in this manner.

"Everybody who uses our software retains full control of their data and can say no to anything they don't want to participate in, but the data is safe to share under HIPAA," LaBonte says, referring to the US health information privacy law. The sheer size of the database allows researchers to assemble a study population that includes almost any set of characteristics.

LaBonte predicts that real-world data will be used increasingly for "pragmatic trials" such as the studies done in 2020 establishing that the malaria drug hydroxychloroquine was not effective against covid-19. "That was a good pragmatic trial candidate because doctors were using it in practice, so people could go through their data, find the patients where it was tried, and then find a matching set of patients who weren't given hydroxychloroquine and call it the control arm," LaBonte says. "So, you've got a trial without ever actually enrolling patients."

An immediate application of external controls is in single-arm trials, in which a company may have only a test arm. Using external data to replace a regular control arm is a higher bar.

"Patients might be more interested because they know everyone will be getting the treatment," Borys says.

Celsion has tried an external control arm to assess a new compound intended for patients with late-stage ovarian cancer. The drug seemed to work well in a test group, but at that point the study did not have a control arm. Medidata Acorn AI compiled an external control arm that was a near-perfect match for the characteristics of the test group, and a comparison of the two groups showed good-enough results to justify Phase 2 trials. "It certainly convinced us that our drug had an effect," Borys says.

The quest for FDA approval

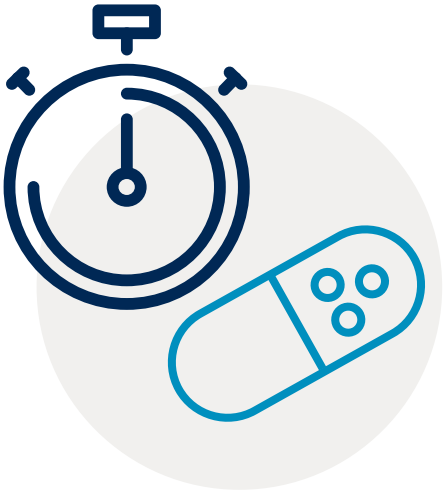
How does the US Food and Drug Administration, the ultimate judge of clinical trial validity, feel about external control arms? Favorable, though appropriately cautious, says Davi, who joined Medidata in 2016 after 22 years as an FDA statistician. She sees one of the best immediate applications in single-arm trials, as with Celsion's ovarian cancer drug, in which a company is evaluating which avenues to pursue and may have only a test arm. In those cases, an external control arm is a vast improvement over the quick-and-dirty comparison method most commonly used: searching the existing medical literature.

For example, a company might compare the results of its treatment with published results on another treatment. If the results are better, it may move into the next phase of developing the new treatment. But the group described in the literature might differ substantially: for example, the company's test group might be mostly women while the subjects from previously published articles are mostly men. Or the test group is from the United States and all the available examples in the literature are from other countries. "You have to have a very large treatment effect for this kind of crude comparison to be convincing, and you're never 100% sure," Davi says.

The published literature can't provide data on the individual patients that were included in a study, but investigators can get that patient-level data using an external control arm created from Medidata's database of past clinical trial patients. They can precisely choose their control patients to match their test patients and get a more scientifically valid comparison of the two groups.

In addition to early phase development, single-arm trials are also used to submit data to the FDA in some circumstances. For example, companies may seek evaluation of a product for accelerated FDA approval in cases in which they think the product represents a dramatic improvement in the standard of care for a disease. The FDA is likely to accept an external control group in such situations because it's arguably better than using summaries from the medical literature or clinical intuition, Davi says.

The next step, using external data to replace a regular control arm, is a much higher bar, Davi says, but even there, progress is being made. Medidata has collaborated on two studies with the nonprofit Friends of Cancer Research to see whether it could build an external control arm that produces results comparable to a standard randomized control arm in clinical trials for treatments of non-small-cell lung cancer and multiple myeloma. In both cases, the studies showed that if the trials had used Medidata Acorn AI's external control arm instead of traditional, randomized controls, the outcomes of the trials would have been the same.



> "Right now, we really don't know if a drug is showing a benefit until we've given it to a few hundred patients. If we have an external control arm, we can get a pretty good signal with a fraction of those numbers."
>
> Arnaub Chatterjee, Senior Vice President for Products, Medidata Acorn AI

In live trials, Davi expects the FDA to be more receptive, at least initially, to a hybrid model, in which the external controls supplement, rather than replace, a standard control arm. Medidata Acorn AI recently helped a customer design a hybrid control arm for a Phase 3 trial – usually the final trial before the FDA approves a drug for general use – for a treatment for recurrent glioblastoma, an aggressive type of brain tumor. Because patients have a poor prognosis and few treatment options, it was important to maximize the number of patients in the trial who receive the experimental treatment. The study design calls for a standard control group that's one third the number of patients that would be necessary for a fully randomized controlled trial. The other two thirds are drawn from historical data. The FDA agreed to consider the results of the hybrid design in the fall of 2020.

A bellwether for effective treatment

External control arms have potential applications after regulatory approval of a product as well, Davi adds. For example, once a treatment is on the market, insurance companies may want to compare it with other treatments for the same condition to determine whether it's cost-effective and whether they will cover it. "The payers may want a comparison to a different study therapy than was used in the randomized controlled trials conducted for regulatory approval. Then the sponsor is faced with having to do a randomized controlled trial against that other therapy," with all the time and expense associated with such trials, Davi says. The sponsor might accomplish the same objective much more efficiently with an external control arm, mining Medidata's database for a matched set of patients treated with the other therapy.

The historical data will also have diminishing validity as care standards change, Borys points out, and older data eventually stops being useful for studying certain things. The performance of a new treatment is generally measured against the current standard of care. The care received by a patient in a historical control arm a few years ago may be substantially different from the care that today's control patient is getting. While the new treatment may perform well by 2015 standards, it needs to perform well by 2021 standards to represent an improvement. A true breakthrough, such as some cutting-edge immunotherapies, can transform the standard of care very quickly and render older patient data irrelevant for many types of studies. "But if you're careful about choosing the endpoints, you still might be able to use the data for some things," Borys adds. And with tens of thousands of clinical trials in progress at any given time, there is always a steady supply of new data.

Beyond the benefits of using external control arms in a clinical trial, Borys also sees potential in being able to track patients after the trial is over: an ability that Medidata can offer. "You can get a better sense of what side effects are really going on in the real world, or how well patients are following up with their treatment, or what are the real costs of treating a patient," he says. "That kind of information is hard to determine in a clinical trial environment, and it could really help us understand how our drugs would fit into a real-world paradigm."

But the greatest value Borys sees currently is in helping to point development of all kinds of treatments in the directions most likely to be successful. "Right now, we really don't know if a drug is showing a benefit until we've given it to a few hundred patients," he says. "If we have an external control arm, we can get a pretty good signal with a fraction of those numbers."

"Clinical trials are better, faster, cheaper with big data" is an executive briefing paper by MIT Technology Review Insights. It is based on research and interviews conducted in March and April 2021. We would like to thank all the participants as well as the sponsor, Medidata. MIT Technology Review Insights has collected and reported on all findings contained in this paper independently, regardless of participation or sponsorship. Jason Sparapani and Laurel Ruma were the editors of this report, and Nicola Crepaldi was the publisher.

About MIT Technology Review Insights

MIT Technology Review Insights is the custom publishing division of MIT Technology Review, the world's longest-running technology magazine, backed by the world's foremost technology institution – producing live events and research on the leading technology and business challenges of the day. Insights conducts qualitative and quantitative research and analysis in the US and abroad and publishes a wide variety of content, including articles, reports, infographics, videos, and podcasts. And through its growing MIT Technology Review Global Panel, Insights has unparalleled access to senior-level executives, innovators, and thought leaders worldwide for surveys and in-depth interviews.

From the sponsor

Medidata is leading the digital transformation of life sciences, creating hope for millions of patients. Medidata helps generate the evidence and insights to help pharmaceutical, biotech, medical device and diagnostics companies, and academic researchers accelerate value, minimize risk, and optimize outcomes. More than one million registered users across 1,700+ customers and partners access the world's most-used platform for clinical development, commercial, and real-world data. Medidata, a Dassault Systèmes company, is headquartered in New York City and has offices around the world to meet the needs of its customers. Discover more at www.medidata.com and follow us @Medidata.

Illustrations
Person icons: kornn. Circular puzzle piece: Arcady. Chart icon: Blan-k. Cardiovascular icon: Stmool. Server icon: Mark stock. Stopwatch icon: NikWB.
All illustrations assembled by Scott Shultz Design, sourced from Shutterstock.

MIT Technology Review Insights

www.technologyreview.com

@techreview @mit_insights

insights@technologyreview.com